UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/610,696 | 07/02/2003 | Amit Srivastava | 02-4024 | 1832 |

7590 02/27/2007

Leonard C. Suchyta
c/o Christian Anderson
Verizon Corporate Services Group Inc.
600 Hidden Ridge, HQE03H01
Irving, TX 75038

| EXAMINER | |
|---|---|
| SIEDLER, DOROTHY S | |
| ART UNIT | PAPER NUMBER |
| 2626 | |

| SHORTENED STATUTORY PERIOD OF RESPONSE | MAIL DATE | DELIVERY MODE |
|---|---|---|
| 3 MONTHS | 02/27/2007 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 10/610,696 | SRIVASTAVA ET AL. | |
| | Examiner | Art Unit | |
| | Dorothy Sarah Siedler | 2626 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *02 July 2003*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-33* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-33* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *02 July 2003* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *9-25-03,12-23-03*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

## DETAILED ACTION

This is the initial response to the application filled on July 2, 2003. Claims 1-33 are pending and are considered below.

### *Claim Rejections - 35 USC § 102*

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

Claims 1,3-5,8,12-15,25,26,28,32, and 33 are rejected under 35 U.S.C. 102(a) as being anticipated by ***Christensen*** ("Punctuation Annotation using Statistical Prosody Models" ISCA Workshop 2001).

As per claims 1,12,32 and 33, ***Christensen*** disclose a linguistic segmentation tool, method, and device comprising: a lexical feature extraction component configured to receive text and generate lexical feature vectors relating to the text (section 1.1 Prosodic and Linguistic clues to structuring speech, last paragraph and section 2.2 Linguistic Information, *textual clues from the words in the text are used to determine punctuation mark classes*) the lexical feature vectors including words from the text and syntactic classes of the words (section 1.1 Prosodic and Linguistic clues to structuring speech, last paragraph and section 2.2 Linguistic Information); an acoustic feature

extraction component configured to receive an audio version of the text and generate acoustic feature vectors relating to the audio version of the text (section 1 Introduction, *prosodic features extracted from the audio data are used*); and a statistical framework component configured to generate linguistic features associated with the text based on the acoustic feature vectors and the lexical feature vectors (section 2.3 Finite State Model Approach, *the words, punctuation mark classes, and prosodic features are combined into a finite state model*).

As per claim 25, ***Christensen*** disclose a method for associating meta-information with a document transcribed from speech, the method comprising: building a language model based on lexical feature vectors extracted from the document, the lexical feature vectors including words and syntactic classifications of the words (section 1.1 Prosodic and Linguistic clues to structuring speech, last paragraph and section 2.2 Linguistic Information); building an acoustic model based on acoustic feature vectors extracted from the speech (section 1 Introduction, *prosodic features extracted from the audio data are used*); and combining outputs of the language model and the acoustic model in a statistical framework that estimates a probability for associating the meta-information with the document (Abstract and section 2.3 Finite State Model Approach, *the words, punctuation mark classes, and prosodic features are combined into a finite state model to determine linguistic meta-data*).

As per claims 3 and 13, ***Christensen*** discloses the linguistic segmentation tool and method of claims 1 and 12, further comprising: a transcription component configured to generate the text based on the audio version of the text (section 1 Introduction, *the ASR system transforms audio into word transcripts*).

As per claim 4, ***Christensen*** discloses the linguistic segmentation tool of claim 1, wherein the statistical framework includes: an acoustic model configured to estimate a probability of an occurrence of the linguistic features based on the acoustic feature vectors (section 2.3, *prosodic features are combined with punctuation classes into a finite state model to determine punctuation*).

As per claim 5, ***Christensen*** discloses the linguistic segmentation tool of claim 4, wherein the statistical framework includes: a language model configured to estimate a probability that one of the lexical feature vectors corresponds to a text boundary (section 2.2 Linguistic information, *words and their corresponding punctuation classes are determined, these classes indicative of commas, periods, questions marks etc. which separate text, specifically sentences and words*).

As per claims 8 and 28, ***Christensen*** discloses the linguistic segmentation tool and method of claims 4 and 25, wherein the acoustic feature vectors are based on prosodic

features including at least one of pause, rate, energy, and pitch (section 2.1 Prosodic Information).

As per claim 14, ***Christensen*** discloses the method of claim 12, further comprising: creating a language model configured to estimate a probability that the lexical features correspond to a word boundary based on the lexical features (section 2.2 Linguistic information, *words and their corresponding punctuation classes are determined, these classes indicative of commas, periods, questions marks etc. which separate text, specifically sentences and words*).

As per claim 15, ***Christensen*** discloses the method of claim 14, further comprising: creating an acoustic model configured to estimate a probability of an occurrence of the linguistic information based on the acoustic features (section 2.3, *prosodic features are combined with punctuation classes into a finite state model to determine punctuation*).

As per claim 26, ***Christensen*** discloses the method of claim 25, wherein the meta-information relates to linguistic features of the document (Abstract, *linguistic meta-data*).

***Claim Rejections - 35 USC § 103***

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 2,6,16,20 and 27 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Christensen***.

As per claims 2 and 20 and 27 ***Christensen*** discloses the linguistic segmentation tool and method of claims 1 and 25, wherein the linguistic features include periods, commas and phrasal boundaries (section 2.2 Linguistic Information). ***Christensen*** does not explicitly disclose the linguistic features including quotation marks and exclamation marks. However, ***Christensen*** does disclose that prosodic and linguistic information combined is used to affectively disambiguate punctuation information in speech (section 1.1 Prosodic and Linguistic clues to Structuring Speech), quotation marks and exclamation marks being common punctuation marks.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to include quotation and exclamation marks as linguistic features in ***Christensen***, since it would enable a system to correctly transcribe text from a spoken

utterance, making the transcript usable for other systems such as information retrieval or speech and natural language understanding.

As per claims 6 and 16, ***Christensen*** discloses the linguistic segmentation tool and method of claims 5 and 15, but does not explicitly disclose wherein the statistical framework includes: a maximum likelihood estimator configured to generate the linguistic features based on the probabilities generated by the acoustic model and the language model. However, using a maximum likelihood estimator is well known in the art, by applicant's own admission (specification page 14).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a maximum likelihood estimator in ***Christensen***, since it is a reliable method to combine the acoustic and lexical (punctuation classes) features, without the need to designate time and resources to develop a new method to combine features.

As per claim 21, ***Christensen*** discloses a computing device for determining linguistic information for words corresponding to a transcribed version of an audio input stream that includes speech that generates lexical features for the words, including a syntactic class associated with at least one of the words (section 1.1 Prosodic and Linguistic clues to structuring speech, last paragraph and section 2.2 Linguistic Information), generates acoustic features for the audio input stream, the acoustic features being

based on at least one of speaker pauses, speaker rate, speaker energy, and speaker pitch (section 1 Introduction, *prosodic features extracted from the audio data are used*), generates the linguistic information based on the lexical features and the acoustic features, and output the generated linguistic information as meta-information embedded in the transcribed version of the audio input stream (section 2.3 Finite State Model Approach and Abstract, *the words, punctuation mark classes, and prosodic features are combined into a finite state model, that information then included as linguistic meta-data for spoken language* ). ***Christensen*** does not explicitly disclose the computing device comprising: a processor; and a computer memory coupled to the processor and containing programming instructions that when executed by the processor, cause the processor to perform the previous steps. However, ***Christensen*** discloses that punctuation annotation systems are often used in conjunction with automatic speech recognition systems (1 Introduction), which are typically performed on a computer with a processor and memory containing software instructions.

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a computing device with a processor and computer memory in ***Christensen***, since a computer can perform computations from program instructions at a speed far greater than a human can manually, therefore saving processing time.

Claims 7,9,10,11,17,18,19,22-24,29,30, and 31 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Cutting***.

***Christensen*** discloses the linguistic segmentation tool and method of claims 1,12,21 and 25 but does not disclose wherein the lexical feature vectors additionally include an identification of a structured speech member of the word, wherein the syntactic classes are indicative of a role of the word in the text, include syntactic classes based on affixes of the words, include syntactic classes based on frequently occurring words, include syntactic classes indicative of the role of the at least one of the words, and wherein the syntactic class is based on affixes of the words. ***Cutting*** discloses lexical feature vectors that include an identification of a structured speech member of the word (page 133 section 1 Desiderata, *words, the lexical features, are assigned parts of speech tags*), wherein the syntactic classes are indicative of a role of the word in the text (page 133 section 1 Desiderata, *the parts of speech tags are used to indicate the linguistic structure of the text*), include syntactic classes based on affixes of the words (page 134 section 2.2 Our Approach, *suffix information is used to predict categories, syntactic classes, for words not in the lexicon*), and include syntactic classes based on frequently occurring words (page 133 section 2.1 Background, *the tags are determined based on models, the models created from probabilities, or frequencies, of each word in a training corpus*). ***Cutting*** discloses that a part-of-speech tagger can be used as input to phrase recognition and grammatical function assignment systems (Abstract).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have lexical features indicative of a structured speech member of a word, include syntactic classes that include the role of a word in text, affixes, and frequency of occurring words in ***Cutting***, since tagging the text with that information is used to determine the linguistic structure of the text, which enables higher-level analysis (page 133 section 1 Desiderata), such as grammatical function assignment and recognizing phrases or other patterns within the text, as indicated in ***Cutting*** (page 133 Abstract and section 1 Desiderata).

## *Conclusion*

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- Suematsu (5,418,716) discloses a system that determines possible grammatical patterns of input sentences.
- Tang (6,718,303) discloses a system for automatically generating punctuation marks in continuous recognition system.
- Chen (6,6067,514) disclose a system for automatically punctuating a speech utterance.
- Mills (7,131,117) discloses a system that analyzes word frequencies from a spoken utterance.

- Nishimura (6,778,958) discloses a system that inserts punctuation marks into a sentence.
- Divay (EP 1,422,692 A2) discloses s system that identifies non-verbalized punctuation in a speech recognition system.
- Oshima (JP 6,1285,570 A) discloses a system that uses speaking intervals, sentence intonation, and parts of speech to determine punctuation mark positions in a sentence structure.
- Shriberg et al ("Can Prosody Aid in th Automatic Processing of Multi-Party Meetings? Evidence from Predicting Punctuation, Disfluencies, and Overlapping Speech" ISCA Tutorial 2001) discloses a system that uses prosody to aid in automatic labeling tasks.
- Beeferman ("CYBERPUNC: A Lightweight Punctuation Annotation System for Speech" IEEE 1998) discloses a system for automatic insertion of intra-sentence punctuation.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dorothy Sarah Siedler whose telephone number is 571-270-1067. The examiner can normally be reached on Mon-Thur 9:30am-5:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DSS

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER